Supporting Information for:

A distinct MaoC-like enoyl-CoA hydratase architecture mediates cholesterol catabolism in *Mycobacterium tuberculosis*

Meng Yang[1], Kip E. Guja[2], Suzanne T. Thomas[1,3], Miguel Garcia-Diaz[2], Nicole S. Sampson[1,*]

[1]Department of Chemistry, Stony Brook University, Stony Brook, NY 11794-3400

[2]Department of Pharmacological Sciences, Stony Brook University, Stony Brook, Y 11794-8651

[3]Current address: Jack H Skirball Center for Chemical Biology and Proteomics, Salk Institute for Biological Studies, La Jolla, CA 92037

*to whom correspondence should be addressed: nicole.sampson@stonybrook.edu, +1-631-632-7952

Supplementary Note.

The asymmetric unit. The two tetramers in one ASU are not identical. Helix α1, α5 and loop I, II in ChsH2^{N} shift positions between the two tetramers. The relative positions of ChsH1 are more conserved compared to ChsH2^{N} (Supplementary Fig. 1a). Similar conformational changes in ChsH2^{N} occur between two heterodimers from the same tetramer (Supplementary Fig. 1b).

Putative role of the DUF35/DUF35_N domain. Electrostatic potential map calculations were undertaken to investigate potential interaction sites with additional proteins or ligands. The electrostatic potential maps clearly revealed that the Coenzyme A moiety resides in a charged environment and the hydrophobic 4-ring system is buried in a neutral environment (Supplementary Fig. 4), as one would expect. The majority of the enzyme complex surface is relatively negative (Supplementary Fig. 4) with one exception; there is a large area of positive potential circled on the surface of ChsH2^{N} (Supplementary Fig. 4b).

ChsH2^{C} is a member of the DUF35/DUF35_N family. SSO2064 is a representative of the DUF35/DUF35_N family for which a structure has been solved (PDB code:3irb). The DUF35/DUF35_N motif is hypothesized to bind and deliver acyl-CoA moieties in acyl-CoA-utilization processes.[1,2] This proposal is consistent with the formation of propionyl-CoA as the end product of β-oxidation of 3-OPC-CoA. It is possible that propionyl-CoA binding to ChsH1-ChsH2 may serve a regulatory function. ChsH2^{N} ends with a long and flexible loop, which would be followed by ChsH2^{C}. The high flexibility of the ChsH2^{C} domain is proposed to deliver generated propionyl-CoAs to other enzymes, for example, enzymes that function in fatty acid biosynthetic pathways. There is a distinctive groove with a hydrophobic and negative potential in SSO2064[1]. This groove matches the hydrophobic 4-ring carbon system and positive potential of

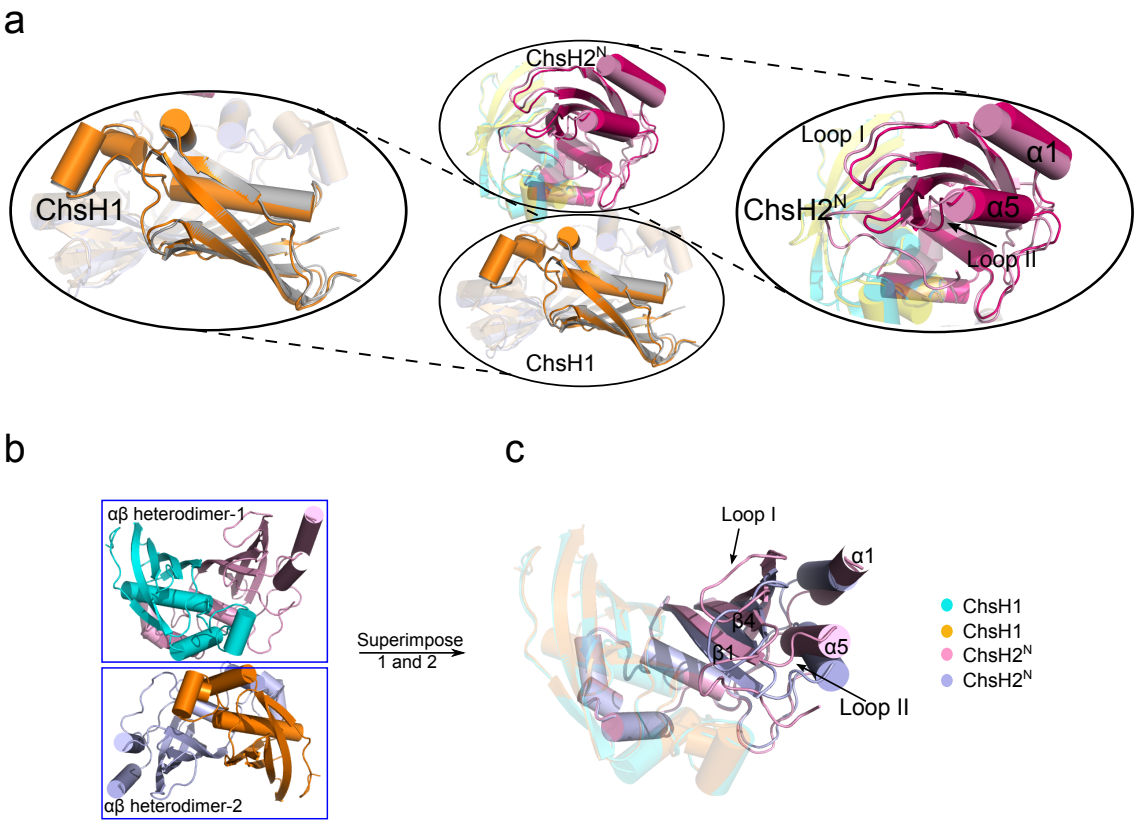
the ChsH2^{N} chain circled in Supplementary Fig. 4b. Therefore, we hypothesize that the ChsH2^{C} domain helps close the substrate tunnel to either positively or negatively regulate catalytic function.

The effects of metals identified in ChsH1-ChsH2N and ChsH1-ChsH2^{N}:3-OPC-CoA structures. Cadmium sites were identified in both the ChsH1-ChsH2^{N} and the ChsH1-ChsH2^{N}:3-OPCC-CoA structure. In the apo-ChsH1-ChsH2^{N} structure, cadmium ions were located around the active site as well as at the ASU packing interface. In the ChsH1-ChsH2^{N}:3-OPC-CoA structure, cadmium ions were identified between the active site and the α,β carbons (17, 20 position) of 3-OPC-CoA that block the interaction of the active site (Asp29, His34) with the substrate analog (Supplementary Fig. 6b). The distance between the active site and the α,β carbons of the substrate analog (8-9 Å) suggests that the cadmium ions bind where the water of hydration is expected to be during catalysis. To confirm whether the metals are introduced by the crystallization process or are required by the enzyme for catalytic activity, ChsH1-ChsH2 was assayed with octenoyl-CoA in the presence of 20 mM EDTA. There was no obvious change in catalytic activity upon addition of EDTA (Supplementary Fig. 6a). However, upon addition of 20 mM or even 2.5 mM cadmium, ChsH1-ChsH2 lost almost all catalytic activity with octenoyl-CoA. The catalytic activity was recovered upon addition of EDTA and cadmium ions simultaneously (Supplementary Fig. 6a), thereby demonstrating that the metal ions from the crystallization inhibit the activity of ChsH1-ChsH2 by blocking the hydration of the substrate. In addition, calcium and cobalt ions from the crystallization process inhibit ChsH1-ChsH2 catalytic activity to some extent at 20 mM. However, no inhibition was observed at 2.5 mM.

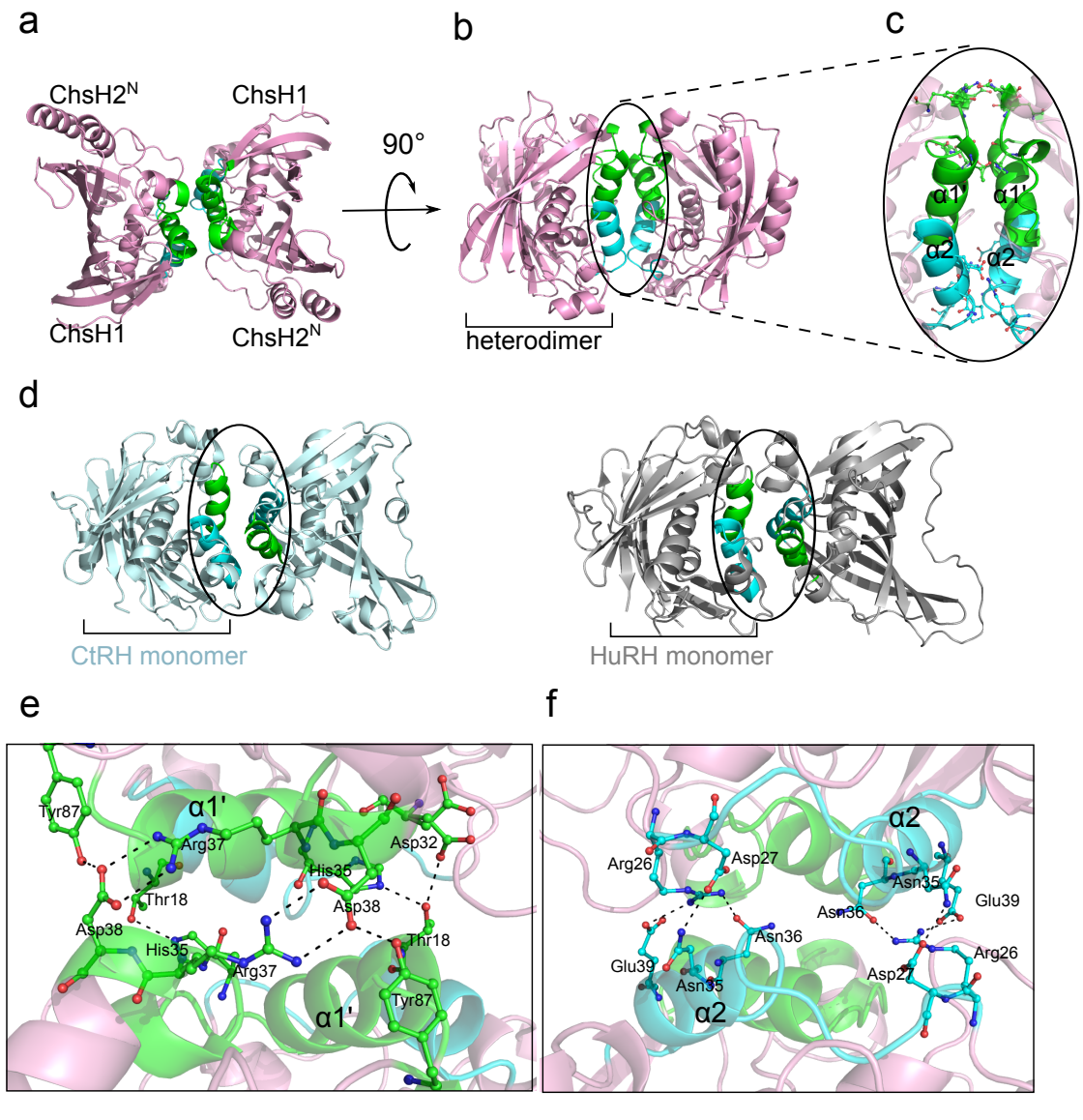
Supplementary Table 1 Comparison of MaoC-like enoyl-CoA hydratases across species.

RH	CHD	ICHD	Number of binding sites	Substrate (aliphatic chain)	Function
AcRH homodimer[3]	Both monomers	None	2	4-6 carbon[4]	Produce (*3R*)-hydroxylacyl-CoA in PHA biosynthesis
CtRH monomer[5]	C-CtRH	N-CtRH	1	10-22 carbon	Peroxisomal β-oxidation of fatty acid to produce (*3R*)-hydroxyacyl-CoAs
HuRH monomer[6]	C-HuRH	N-HuRH	1	Methyl-branched;	Breakdown of very-long or methyl-branched fatty acids; bile acid synthesis
ChsH1-ChsH2 heterodimer	ChsH1	ChsH2	1	steroid 3-carbon side chain	β-oxidation of cholesterol in *Mtb*

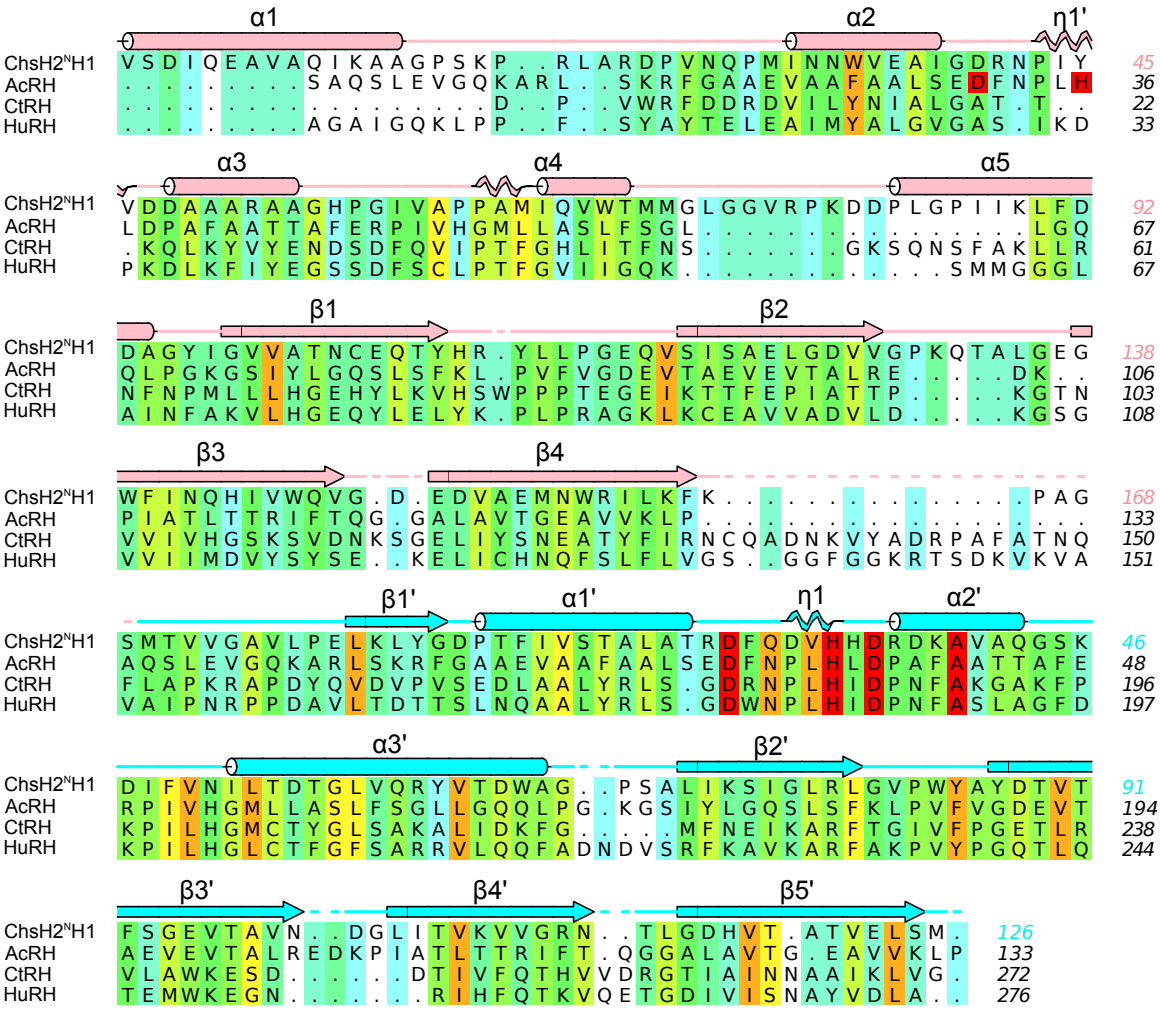
CHD: complete hot-dog fold; ICHD: incomplete hot-dog fold
AcRH: *Aeromonas caviae* MaoC-like enoyl-CoA hydratase
CtRH: *Candida tropicalis* MaoC-like enoyl-CoA hydratase
HuRH: *Homo sapiens* MaoC-like enoyl-CoA hydratase
C-CtRH; C-HuRH: C-terminal domain of CtRH; C-terminal domain of HuRH
N-CtRH; N-HuRH: N-terminal domain of CtRH; N-terminal domain of HuRH


a
ChsH2N
ChsH1
Loop I
α1
ChsH2N
α5
Loop II
ChsH1
b
αβ heterodimer-1
αβ heterodimer-2
Superimpose
1 and 2
c
Loop I
α1
β4
β1
α5
Loop II
ChsH1
ChsH1
ChsH2N
ChsH2N

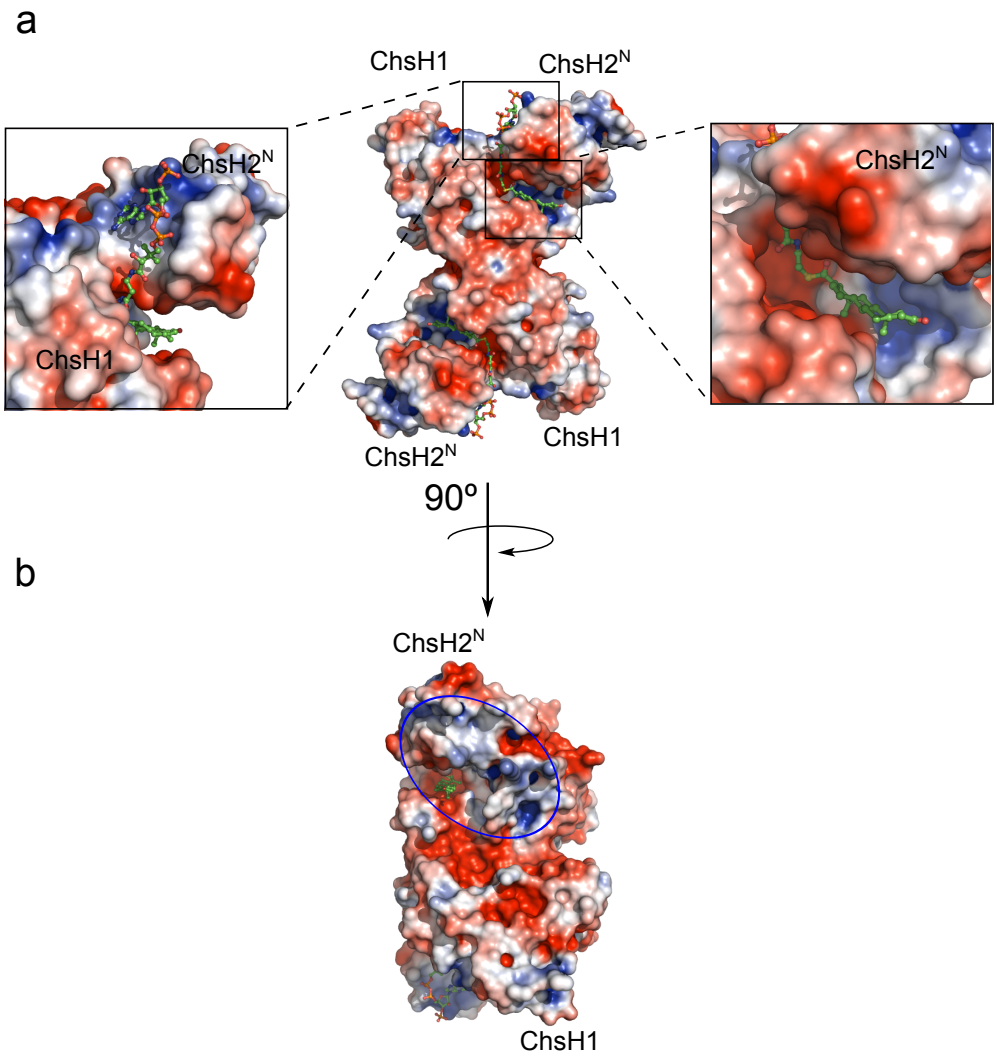
Supplementary Figure 1 Structures in the Asymmetric unit. (**a**) Superimposition of two heterotetramer structures from one asymmetric unit. One set of overlayed ChsH1 chains (orange and gray) and ChsH2^{N} chains (magenta and pink) are labeled. The other set of overlayed ChsH1 and ChsH2^{N} chains is rendered transparent. The conformation of ChsH1 is conserved without any obvious variation between heterotetramers. The variations of secondary structures in ChsH2^{N} are labeled. (b) Superimposition of two heterodimers from the same tetramer. Overlayed ChsH1 chains (cyan and orange) and ChsH2^{N} chains (pink and purple) are labeled. The ChsH1 chains are more conserved compared to ChsH2^{N} chains. The overlayed ChsH1 chains are rendered transparent in (c). The variations in secondary structures in ChsH2^{N} are labeled.


a
b
c
ChsH2N
ChsH1
ChsH1
ChsH2N
90°
heterodimer
α1
α1'
α2
α2
d
CtRH monomer
HuRH monomer
e
f
α1'
Tyr87
Arg37
Asp32
His35
Thr18
Asp38
Asp38
His35
Arg37
Thr18
Tyr87
α1'
α2
Arg26
Asp27
Asn35
Glu39
Asn36
Asn36
Glu39
Arg26
Asp27
Asn35
α2

Supplementary Figure 2 Dimer-dimer interface in ChsH1-ChsH2^N heterotetramer. (**a**) Overall structure of ChsH1-ChsH2^N tetramer. The chain identities are labeled. All the chains are colored in pink except the secondary structures in the dimer-dimer interface. Helices α1' from ChsH1 (green) and helices α2 from ChsH2^N (cyan) build up the heterodimer-heterodimer interface. (**b**) The structure from (a) is rotated 90° about the X-axis. The secondary structure elements in the interface are labeled. (**c**) Magnification of the dimer-dimer interface. Residues that are responsible for the dimer-dimer interactions are shown as sticks. Interacting residues from ChsH1 chains and ChsH2^N chains are colored in green and cyan, respectively. (**d**) The monomer-monomer interface in AcRH and HuRH. Helices from the C-terminal of AcRH or HuRH (green) and helices from the N-terminal of AcRH or HuRH (cyan), which are counterparts of α1' and α2, respectively, constitute the monomer-monomer interface[5,6]. (**e**) Tyr87 and Thr18 from one ChsH1 chain form hydrogen bonds with Asp38 and His35, respectively, from the second ChsH1 chain. One conformation of Asp32 in the first ChsH1 chain forms hydrogen bonds with Thr18 of the second ChsH1 chain. (**f**) Arg26 and Asp27 from one ChsH2^N chain hydrogen bond with Glu39 and Asn36, and Asn35, respectively from the other ChsH2^N chain.


α1
α2
η1'
ChsH2NH1 VSDIQEAVAQIKAAGPSKP..RLARDPVNQPMINNWVEAIGDRNPIY 45
AcRH .....SAQSLEVGQKARL..SKRFGAAEVAAFAALSEDFNPLH 36
CtRH ..........D..P..VWRFDDRDVILYNIALGAT.T.. 22
HuRH .....AGAIGQKLPP..F..SYAYTELEAIMYALGVGAS.IKD 33
α3
α4
α5
ChsH2NH1 VDDAAARAAGHPGIVAPPAMIQVWTMMGLGGVRPKDDPLGPIIKLFD 92
AcRH LDPAFAATTAFERPIVHGMLLASLFSGL.....LGQ 67
CtRH .KQLKYVYENDSDFQVIPTFGHLITFNS.....GKSQNSFAKLLR 61
HuRH PKDLKFIYEGSSDFSCLPTFGVIIGQK.....SMMGGGL 67
β1
β2
ChsH2NH1 DAGYIGVVATNCEQTYHR.YLLPGEQVSISAELGDVVGPKQTALGEG 138
AcRH QLPGKGSIYLGQSLSFKL.PVFVGDEVTAEVEVTALRE.....DK.. 106
CtRH NFNPMLLLHGEHYLKVHSWPPPTEGEIKTTFEPIATTP.....KGTN 103
HuRH AINFAKVLHGEQYLELYK.PLPRAGKLKCEAVVADVLD.....KGSG 108
β3
β4
ChsH2NH1 WFINQHIVWQVG.D.EDVAEMNWRILKFK.....PAG 168
AcRH PIATLTTRIFTQG.GALAVTGEAVVKLP..... 133
CtRH VVIVHGSKSVDNKSGELIYSNEATYFIRNCQADNKVYADRPAFATNQ 150
HuRH VVIIMDVYSYSE..KELICHNQFSLFLVGS..GGFGGKRTSDKVKVA 151
β1'
α1'
η1
α2'
ChsH2NH1 SMTVVGAVLPELKLYGDPTFIVSTALATRDFQDVHHDRDKAVAQGSK 46
AcRH AQSLEVGQKARLSKRFGAAEVAAFAALSEDFNPLHLDPAFAATTAFE 48
CtRH FLAPKRAPDYQVDVPVSEDLAALYRLS.GDRNPLHIDPNFAKGAKFP 196
HuRH VAIPNRPPDAVLTDTTSLNQAALYRLS.GDWNPLHIDPNFASLAGFD 197
α3'
β2'
ChsH2NH1 DIFVNILTDTGLVQRYVTDWAG..PSALIKSIGLRLGVPWYAYDTVT 91
AcRH RPIVHGMLLASLFSGLLGQQLPG.KGSIYLGQSLSFKLPVFVGDEVT 194
CtRH KPILHGMCTYGLSAKALIDKFG.....MFNEIKARFTGIVFPGETLR 238
HuRH KPILHGLCTFGFSARRVLQQFADNDVSRFKAVKARFAKPVYPGQTLQ 244
β3'
β4'
β5'
ChsH2NH1 FSGEVTAVN..DGLITVKVVGRN..TLGDHVT.ATVELSM. 126
AcRH AEVEVTALREDKPIATLTTRIFT.QGGALAVTG.EAVVKLP 133
CtRH VLAWKESD.....DTIVFQTHVVDRGTIAINNAAIKLVG. 272
HuRH TEMWKEGN.....RIHFQTKVQETGDIVISNAYVDLA.. 276

Supplementary Figure 3 Secondary structure sequence alignment of MaoC-like enoyl-CoA hydratases from different species. (**a**) Structural sequence alignment between ChsH1-ChsH2^{N} dimer, AcRH homodimer, CtRH monomer (PDB code: 1PN2), and HuRH monomer (PDB code: 1S9C). The sequence of ChsH2^{N} (1-178) and the sequence of ChsH1 (1-127) overlay with the N-terminal and C-terminal sequences of the monomers, respectively. The heterodimer sequence is referred to as ChsH2^{N}H1. Secondary structure elements are labeled and color coded following Fig. 3d and 3e. The residues are colored based on their conservation. The calculated RMS is 1.969 Å with 969 atoms aligned after the superimposition between AcRH monomer and ChsH1-ChsH2^{N} dimer. The calculated RMS is 2.331 Å with 818 atoms aligned after the superimposition between CtRH monomer and ChsH1-ChsH2^{N} dimer. The calculated RMS is 1.826 Å with 1009 atoms aligned after the superimposition between HuRH monomer and ChsH1-ChsH2^{N} dimer. The protein sequence of ChsH2^{N}H1 dimer shares 47.4% sequence identity with CtRH monomer and shares 47.1% with HuRH monomer.

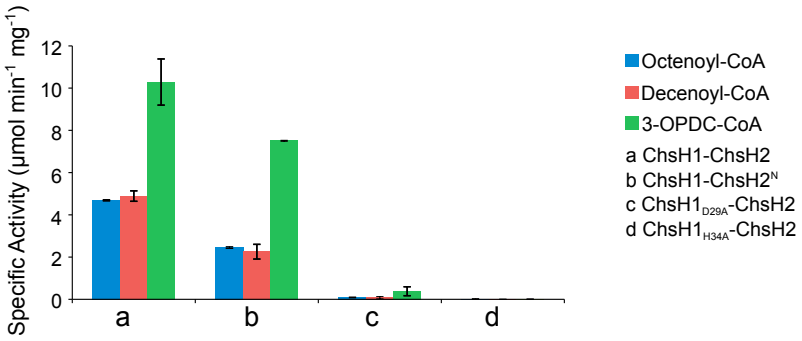

a
ChsH1
ChsH2N
ChsH2N
ChsH1
ChsH2N
ChsH1
ChsH2N
90º
b
ChsH2N
ChsH1

Supplementary Figure 4. Electrostatic surface potential map of ChsH1-ChsH2^{N}:3-OPC-CoA. The electrostatic maps were generated using APBS in Pymol, positive potential is blue (+3 kTe-1) and negative potential is red (-3kTe-1). 3-OPC-CoA is colored by atom.

a

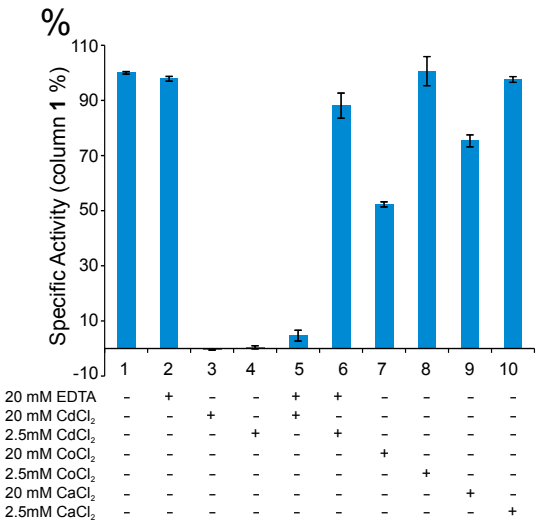
Aeromonas caviea	(I)	23-VAAFAALSEDFNPLHLDPAFA-43
Homo sapiens	(II)	184-QAALYRLSGDWNPLHIDPNFA-204
Candida tropicalis	(II)	173-LAALYRLSGDRNPLHIDPNFA-193
RAM_12770	(III)	21-VVSTALATRDFQDVHHDRDSA-41
SSMG_02035	(III)	26-VISTALATRDFQDVHHDRDAA-46
Gbro_0886	(III)	27-VVSTALATRDFQDVHHDRDLA-47
RHA1_ro04487	(III)	25-VVSAALATRDFQDVHHDRDLA-45
AS9A_0909	(III)	27-IISTAIATRDFEDVHHDRDAA-47
MAV_0620	(III)	24-IISTALATRDFQDVHHDRDKA-44
Mb3571c	(III)	20-IVSTALATRDFQDVHHDRDKA-40
Rv3541c(ChsH1)	(III)	20-IVSTALATRDFQDVHHDRDKA-40
Swit_3320	(IV)	22-IVAAAIATNDYEDVHHDKAAA-42
IMCC3088_1454	(IV)	21-IVAGAIASRDYQNVHHDKDAA-41
Spea_2140	(IV)	35-ITSGAIATRDFFPGHHDKDAA-55
AMIS19050	(V)	301-IISTALATRDFQDVHHDRDAA-321
FrRUN1fDRAFT_2215	(V)	434-IVSGAIATRDFQPVHHDRELA-454

b

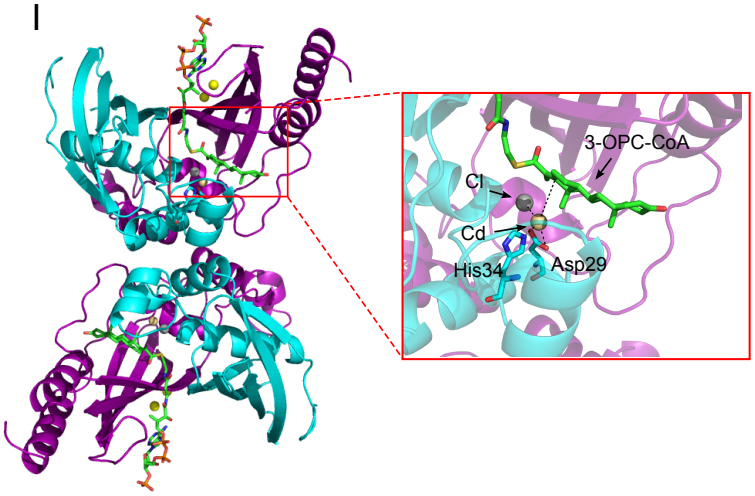


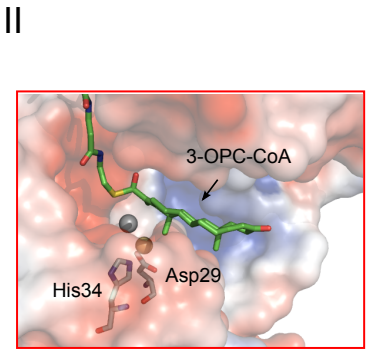
Supplementary Figure 5 Protein sequence alignment of ChsH1 and its homologs across species and specific activity data for ChsH1$_{D29A}$-ChsH2 and ChsH1$_{H34A}$-ChsH2. **(a)** Active site housing segment sequence alignment of ChsH1 against other MaoC-like enoyl-CoA hydratases across species. All the organisms included in the alignment match the organisms in the phylogenetic tree in Fig. 7. The operonic organizations of genes encoding MaoC-like enoyl-CoA hydratases were labeled following Fig. 7. All the aligned residues are colored by similarity (ClustalW). The highly conserved putative active siteresidues are highlighted yellow. The sequence accession numbers from NCBI for each organism are *Aeromonas Caviea* (O32472.1), *Homo sapiens* (PDB: 1S9C), *Candida tropicalis* (PDB:1PN2), RAM_12770 (*Amycolatopsis mediterranei S699*; YP_005530501.1), SSMG_02035 (*Streptomyces sp. AA4*; ZP_07277995.1), Gbro_0886 (*Gordonia bronchialis DSM 43247*; YP_003272092.1), RHA1_ro04487 (*Rhodococcus jostii RHA1*; YP_704431.1), AS9A_0909 (*Amycolicicoccus subflavus DQS3-9A1*; YP_004492161.1), MAV_0620 (*Mycobacterium avium 104*; YP_879900.1), Mb3571c (*Mycobacterium bovis AF2122/97*; NP_857210.1), ChsH1 (*Mycobacterium tuberculosis H37Rv*; NP_218058.1), Swit_3320 (*Sphingomonas wittichii RW1*; YP_001263804.1), IMCC3088_1454 (*Gamma proteobacterium IMCC3088*; ZP_08270994.1), Spea_2140 (*Shewanella pealeana ATCC 700345*; YP_001501995.1), AMIS19050 (*Actinoplanes missouriensis 431*; YP_005461641.1), FrEUN1fDRAFT_2215 (*Frankia sp. EUN1f*; ZP_06412519.1). **(b)** Specific activity data for ChsH1-ChsH2, ChsH1-ChsH2^{N}, ChsH1$_{D29A}$-ChsH2 and ChsH1$_{H34A}$-ChsH2 assayed with octenoyl-CoA, decenoyl-CoA and 3-OPC-CoA.

a


%
Specific Activity (column 1 %)
110
90
70
50
30
10
-10
1 2 3 4 5 6 7 8 9 10
20 mM EDTA - + - - + + - - - -
20 mM CdCl2 - - + - + - - - - -
2.5mM CdCl2 - - - + - + - - - -
20 mM CoCl2 - - - - - - + - - -
2.5mM CoCl2 - - - - - - - + - -
20 mM CaCl2 - - - - - - - - + -
2.5mM CaCl2 - - - - - - - - - +

b


I
3-OPC-CoA
Cl
Cd
His34
Asp29


II
3-OPC-CoA
His34
Asp29

Supplementary Figure 6 The effects of metals used for crystallization on catalytic activity. (a) The effects of metal ions on the activity of ChsH1-ChsH2. (b) I, 6 Cd^{2+} ions and 2 Cl^{-1} are identified in one of the ChsH1-ChsH2^{N} heterotetramers. One Cd^{2+} is located in the middle of the active site near the C17-C20 position of 3-OPC-CoA. Cl^{-1} coordinates with Cd^{2+}. II, electrostatic map shows an obvious negative pocket in which Cd^{2+} is located.

Supplementary References

1. Krishna, S. S., Aravind, L., Bakolitsa, C., Caruthers, J., Carlton, D., Miller, M. D., Abdubek, P., Astakhova, T., Axelrod, H. L., Chiu, H. J., Clayton, T., Deller, M. C., Duan, L., Feuerhelm, J., Grant, J. C., Han, G. W., Jaroszewski, L., Jin, K. K., Klock, H. E., Knuth, M. W., Kumar, A., Marciano, D., McMullan, D., Morse, A. T., Nigoghossian, E., Okach, L., Reyes, R., Rife, C. L., van den Bedem, H., Weekes, D., Xu, Q., Hodgson, K. O., Wooley, J., Elsliger, M. A., Deacon, A. M., Godzik, A., Lesley, S. A., and Wilson, I. A. (2010) The structure of SSO2064, the first representative of Pfam family PF01796, reveals a novel two-domain zinc-ribbon OB-fold architecture with a potential acyl-CoA-binding role., *Acta Crystallogr. Sect. F Struct. Biol. Cryst. Commun. 66*, 1160-1166.
2. Krishna, S. S., Weekes, D., Bakolitsa, C., Elsliger, M. A., Wilson, I. A., Godzik, A., and Wooley, J. (2010) TOPSAN: use of a collaborative environment for annotating, analyzing and disseminating data on JCSG and PSI structures., *Acta Crystallogr. Sect. F Struct. Biol. Cryst. Commun. 66*, 1143-1147.
3. Hisano, T., Tsuge, T., Fukui, T., Iwata, T., Miki, K., and Doi, Y. (2003) Crystal structure of the (*R*)-specific enoyl-CoA hydratase from *Aeromonas caviae* involved in polyhydroxyalkanoate biosynthesis, *J. Biol. Chem. 278*, 617-624.
4. Fukui, T., Shiomi, N., and Doi, Y. (1998) Expression and characterization of (*R*)-specific enoyl coenzyme A hydratase involved in polyhydroxyalkanoate biosynthesis by *Aeromonas caviae*, *J. Bacteriol. 180*, 667-673.
5. Koski, M. K., Haapalainen, A. M., Hiltunen, J. K., and Glumoff, T. (2004) A Two-domain structure of one subunit explains unique features of eukaryotic hydratase 2, *J. Biol. Chem. 279*, 24666-24672.
6. Koski, K. M., Haapalainen, A. M., Hiltunen, J. K., and Glumoff, T. (2005) Crystal structure of 2-enoyl-CoA hydratase 2 from human peroxisomal multifunctional enzyme type 2, *J. Mol. Biol. 345*, 1157-1169.